Component of Statistics Canada Catalogue no. 82-003-X
Health Reports

Article

Evaluation of the coverage of linked Canadian Community Health Survey and hospital inpatient records

by Michelle Rotermann

January 2009

Statistics Canada Statistique Canada

Canada

Evaluation of the coverage of linked Canadian Community Health Survey and hospital inpatient records

by Michelle Rotermann

Abstract

Background

Evaluation of the coverage that results from linking routinely collected administrative hospital data with survey data is an important preliminary to undertaking analyses based on the linked file.

Data and methods

To evaluate the coverage of the linkage between data from cycle 1.1 of the Canadian Community Health Survey (CCHS) and in-patient hospital data (Health Person-Oriented Information or HPOI), the number of people admitted to hospital according to HPOI was compared with the weighted estimate for CCHS respondents who were successfully linked to HPOI. Differences between HPOI and the linked and weighted CCHS estimate indicated linkage failure and/or undercoverage.

Results

According to HPOI, from September 2000 through November 2001, 1,572,343 people (outside Quebec) aged 12 or older were hospitalized. Weighted estimates from the linked CCHS, adjusted for agreement to link and plausible health number, were 7.7% lower. Coverage rates were similar for males and females. Provincial rates did not differ from those for the rest of Canada, although differences were apparent for the territories. Coverage rates were significantly lower among people aged 75 or older than among those aged 12 to 74.

Keywords

coverage, data collection, databases, health surveys, hospital records, inpatients, medical records, probabilistic linkage

Author

Michelle Rotermann (1-613-951-3166, Michelle.Roterman@statcan.gc.ca) is with the Health Information and Research Division at Statistics Canada, Ottawa, Ontario, K1A 0T6.

Record linkage is used in health studies to obtain more complete information, to fill gaps in existing datasets, and/or to improve data quality.[1,2] For instance, prospective death clearance of survey respondents, study cohorts or administrative data sources, such as inpatient hospital records, have made it possible to study associations between death and factors such as lifestyle, occupation, treatment modalities, patient histories and geography.[3-20]

Similarly, linking birth and stillbirth records with death registrations and/or hospitalization data has enabled the study of maternal, fetal and infant morbidity and mortality by maternal and infant characteristics.[21-24] Record linkage has also been used to validate self-reported information,[25,26] describe the characteristics of unmatched records,[27] assess the comparability or quality of data files generated using probabilistic and deterministic linkage approaches,[28] reduce under-ascertainment of disease prevalence,[29] and monitor health system performance.[30,31] In the absence of disease registries, record linkage is a cost-effective and efficient way to monitor disease incidence and prevalence.[32-35]

This study was motivated by the need to assess the coverage of the linkage between the Canadian Community Health Survey (CCHS) and Health Person-Oriented Information (HPOI), *an administrative database of hospital records*. Initial research on the rate of linkage between the CCHS and HPOI estimated the proportion of CCHS respondents who had been hospitalized during the 1994/1995 to 2004/2005 period, but coverage has yet to be assessed.[36] Evaluation of the coverage is essential if the linked file is to be used for epidemiologic research. It is important to know if findings will be biased, that is, if survey respondents with certain characteristics are more likely than others to have been linked.

HPOI and the CCHS are complementary sources of data. HPOI does not have information about non-medical determinants of health, such as socio-economic and lifestyle factors.

For example, hospital records do not contain information about smoking status or body mass index (BMI), two important risk factors. The CCHS, by contrast, is a rich source of information about health status and determinants of health, but lacks the detail needed to study hospitalization. Combining HPOI with the CCHS reduces many of the limitations of each source, and thereby facilitates a more complete understanding of what brings Canadians in contact with the health care system and how they fare within the system.

The two main objectives of this study were to: 1) evaluate the coverage of the linked CCHS and HPOI by calculating coverage rates; and 2) identify characteristics of CCHS cycle 1.1 respondents who were less likely to be in the linked file.

Methods

Data sources

Canadian Community Health Survey

The Canadian Community Health Survey is a cross-sectional survey that collects information about health status, health care use and health determinants. It covers the household population aged 12 or older in the provinces and territories, except members of the regular Forces and residents of institutions, Indian reserves and other Aboriginal settlements, and some remote areas. The rate of coverage is in the 98% range in the provinces, 97% in the Northwest Territories, 90% in the Yukon, and 71% in Nunavut.

Data for cycle 1.1 were collected from September 1, 2000 through November 3, 2001 from a sample of 131,535 people; the response rate was 84.7%. All CCHS information, including provincial health care numbers (HNs) and postal codes, is self-reported by respondents, and the extent of error in these variables is unknown. However, data capture applications used by interviewers contain features that check for inconsistent answers, out-of-range responses or invalid alpha-numeric sequences. More information about the CCHS is available in a published report.[37]

CCHS respondents were asked for permission to link information collected during the interview with their provincial health information, including past and continuing use of services such as hospitals, clinics, doctor's offices or other services provided by the province; 91% of respondents gave permission. The sample used for this study consists of 72,354 (66.5%) respondents aged 12 or older in all provinces and territories except Quebec, who agreed to link and provided a valid health number (HN) (Appendix A). Quebec HPOI records cannot be linked to CCHS records because the Quebec hospital records provided to Statistics Canada contain scrambled HNs, no date of birth and incomplete postal codes.

Survey weights were used so that estimates produced from the CCHS data were representative of the target population, not just the sample itself. The survey weight is the number of people in the population represented by each respondent. Survey weights reflect the differing probabilities of selection and response. Each record is, therefore, weighted by the inverse of the probability of selecting the person and getting a response from him or her.[38] Additional survey weights are required for record linkage because not all respondents agree to link and not all those who agree to link, provide a valid HN. For this study, survey weights, adjusted for agreement to link and provision of a valid HN, were calculated.

Statistics Canada does not have access to provincial health insurance databases against which the HNs provided by CCHS respondents could be verified. Instead, all provinces and territories provide check-digit formulas that are used to verify that the HNs are at least plausible. Although check-digits are not a substitute for databases that contain first and last names, birth dates, addresses and HNs, they can detect accidental transcription errors, such as the inversion of two numbers, and offer a simple method of distinguishing meaningful numbers from strings of random digits.

Hospital data

The Health Person-Oriented Information (HPOI) database, maintained by Statistics Canada, contains information about inpatient hospital separations (discharges and in-hospital deaths) from virtually all acute-care and some psychiatric, chronic and rehabilitative hospitals.

HPOI is a person-level dataset derived from discharge records (which can reflect multiple discharges of the same person) in the Hospital Morbidity Database (HMDB). Sequential person-level HPOI records can be used to construct each patient's hospitalization history. During the linkage process, records belonging to the same individual are identified from the patient's HN and demographic and diagnosis/intervention information (for example, sex, birth date, sex-specific procedures).[39]

Hospital records pertaining to the past fiscal year are added to HPOI annually. With each additional year of data, the entire HPOI process is rerun to ensure internal consistency of the demographic information at the person-level for patients with multiple hospital discharges.

Reabstraction studies, which validate the accuracy of hospital records, have found that the non-medical administrative data elements (essential for record linkage) are of high quality. For example, 99% of a random sample of discharge records for hospital stays from September through November 2000 had correct HNs, and 91% of postal codes were error-free.[40]

Statistics Canada has hospital data with HNs for all provinces (except Quebec) and the Northwest Territories from fiscal year 1994/1995 onwards;

data for 1992/1993 and 1993/1994 are available for some provinces.

While the HPOI database includes the vast majority of records from HMDB, about 3% of records for patients aged 12 or older (the target population of this study) were excluded because of missing or invalid HNs.[39]

From September 1, 2000 through November 3, 2001, there were 2.3 million discharges of 1,624,972 people aged 12 or older from acute-care hospitals outside Quebec. Discharges from non-acute hospitals were excluded from this study because coverage of such hospitals is inconsistent across provinces.

The target populations of the Canadian Community Health Survey (CCHS) and HPOI differ somewhat. The CCHS excludes full-time members of the Canadian Forces and residents of Indian Reserves, of institutions (for instance, nursing homes and prisons) and of some remote areas. HPOI is a census and, therefore, these groups are included among hospitalizations. In an effort to match the target populations of the CCHS and HPOI more closely, hospitalizations that could be identified as pertaining to the on-reserve or the institutionalized population were removed from this analysis.

The on-reserve population is a derived census variable created by identifying census sub-division (CSD) type according to criteria established by Indian and Northern Affairs Canada (INAC), as well as selected CSDs that correspond to northern communities in Saskatchewan, the Northwest Territories, and the Yukon.[41] The postal code conversion file (PCCF+)[42] and a list of facilities used by the Residential Care Facility survey[43] were used to identify institutional residents. Hospitalizations pertaining to 31,330 residents of Reserves and associated lands were removed from HPOI, as were hospitalizations of 21,299 residents of institutions. Removal of these 52,629 records, which amounted to about 3% of the HPOI patients hospitalized during the study period, brought the population covered by HPOI more in line with the CCHS target population.

Analytical techniques

Probabilistic record linkage

Probabilistic record linkage was used to identify CCHS respondents who were hospitalized. The linkage between the CCHS and HPOI was done with Generalized Record Linkage software (GRLS) developed at Statistics Canada. The two data sources contain many variables, but only a few fields appear in both and are distinct enough to be useful in matching for linkage. A CCHS respondent was considered to have been hospitalized if a record containing an HN and/or similar demographic characteristics (for example, birth date, sex, postal code) and an admission date to an acute-care facility between September 1, 2000 and November 3, 2001 was found in HPOI.

Probabilistic linkage does not require complete agreement on the matching variables. Rather, the quality of the match between pairs of records is rated with algorithms that evaluate the likelihood of a correct match[1,44] (Figure 1). Points were given or subtracted depending on the similarity of the values between fields. For instance, high positive scores were assigned if the HNs were identical and the issuing province of the HN matched; if the values were similar but not exact, a lower positive score was assigned, reflecting partial agreement; if the values

Figure 1
Example of how pairs of Canadian Community Health Survey (CCHS) and Health Person-oriented Information (HPOI) records were assessed and scored using Generalized Record Linkage Software (GRLS)

ID	Province	Birthdate	Postal code	Health number (HN)	Sex	Score calculated by GRLS	Match?	Commentary
A	Ontario	11/06/1964	L9Y3B9	3512345678	Female	447	yes	All fields match
1	Ontario	11/06/1964	L9Y3B9	3512345678	Female			
F	Manitoba	24/07/1927	R0A0T0	55667788	Male	-308	no	Nothing matches
1	Ontario	11/06/1964	L9Y3B9	3512345678	Female			
B	Manitoba	21/05/1945	R0A0T0	missing	Male	268	yes	HN missing; everything else matches
7	Manitoba	21/05/1945	R0A0T0	4624252627	Male			
B	Manitoba	21/05/1945	R0A0T0	missing	Male	-244	no	HN missing; nothing else matches
1	Ontario	11/06/1964	L9Y3B9	3512345678	Female			

☐ CCHS record ☐ HPOI record

on the two records were totally different, points were subtracted.

The number of points assigned to each pair of linking variables reflected their importance as matching variables, which typically was related to uniqueness. For example, because there are only two possible values for the sex of the respondent/patient, matches on this field scored fewer points than if the postal codes or HNs matched.

Total linkage weights for each pair of CCHS-HPOI records were calculated by summing the scores assigned to each pair of linking variables. The higher the total linkage weight, the more likely the two records pertained to the same individual. Total linkage weights ideally form a bi-modal distribution. When pairs of records scored above the selected threshold, they were accepted as "true" matches; pairs below the threshold were rejected. To eliminate the need for manual review, the cut-off points chosen for this study were identical, which meant that each pair of records could have only one of two values: match or non-match.

Results

To evaluate the coverage of the linkage between cycle 1.1 of the CCHS and HPOI, the number of people admitted to hospital according to each data source was compared. Survey weights, adjusted for agreement to link and HN validity, were applied to the records of CCHS respondents for whom records were also found in the HPOI database. The HPOI count of hospitalizations was regarded as the standard. The coverage rate was calculated by dividing the weighted estimates of CCHS respondents who successfully linked to HPOI by HPOI counts, minus records identified as pertaining to residents of Indian Reserves or associated lands or of institutions and then multiplying by 100. Differences between the HPOI counts and the weighted estimates from the CCHS were examined. Standard errors and 95% confidence intervals were calculated for the coverage rates using the bootstrap technique. Statistical significance was tested using the t-test ($p<0.05$).[45,46]

Table 1
Number hospitalized in acute-care hospitals and coverage rates, September 1, 2000 through November 3, 2001, by selected characteristics and data source, population aged 12 or older, Canada excluding Quebec

	Health Person-Oriented Information (HPOI)	Canadian Community Health Survey (CCHS)		CCHS/HPOI Coverage rate		
					95% confidence interval	
	Number	Unweighted number	Weighted number	%	from	to
Total	1,572,343	6,785	1,451,272	92.3	88.9	95.7
Province/Territories						
Newfoundland and Labrador	41,394	272	40,445	97.7	83.6	111.8
Prince Edward Island	11,784	237	11,061	93.9	79.6	108.1
Nova Scotia	67,226	348	60,419	89.9	78.0	101.7
New Brunswick	67,542	423	62,203	92.1	81.7	102.5
Ontario	753,970	2,230	694,463	92.1	86.6	97.6
Manitoba	82,386	567	69,739	84.6	73.6	95.7
Saskatchewan	82,778	659	78,664	95.0	86.4	103.7
Alberta	202,498	863	186,301	92.0	83.3	100.7
British Columbia	258,883	1,062	241,647	93.3	85.3	101.3
Territories	3,882	124	6,331	163.1*	139.3	186.9
Sex						
Female†	971,087	4,343	904,318	93.1	88.8	97.5
Male	601,249	2,442	546,955	91.0	85.4	96.5
Age group						
12 to 24†	173,009	711	165,093	95.4	81.5	101.3
25 to 34	276,150	1,041	270,164	97.8	85.0	105.8
35 to 44	210,848	830	209,447	99.3	88.3	110.4
45 to 54	183,878	773	174,194	94.7	84.9	104.6
55 to 64	181,041	854	169,235	93.5	83.2	103.7
65 to 74	227,410	1,090	219,259	96.4	86.9	106.0
12 to 74	1,252,336	5,299	1,207,392	96.4	92.4	100.4
75 or older	320,007	1,486	243,881	76.2*	70.2	82.2

† reference category
* significantly different from reference category ($p<0.05$); for provincial comparison, significantly different from rest of Canada, for example, Ontario compared with Canada minus Ontario
Source: 2000/2001 Canadian Community Health Survey; Health Person-oriented Information, 2000/2001 to 2001/2002

According to HPOI, from September 1, 2000 through November 3, 2001, 1,572,343 people were admitted to an acute-care hospital (excluding Quebec) (Table 1). Weighted estimates from the CCHS, adjusted for agreement to link and valid HN, were 7.7% lower (1,451,272).

Coverage rates were similar for males and females (91.0% and 93.1%). Provincial rates did not differ significantly from the rate for the rest of Canada. However, based on the CCHS, the estimated number of residents of the territories who were hospitalized was considerably higher than the HPOI number. As a result, the coverage rate for the territories exceeded 100%.

Coverage rates for most age groups were similar. The exception was seniors aged 75 or older whose rate (76.2%) was significantly lower than that of people aged 12 to 74 (96.4%).

Discussion

The significantly lower coverage rate for seniors aged 75 or older was anticipated because the two data sources did not pertain to exactly the same populations. The CCHS excludes residents of institutions, but they are included in the hospital data (HPOI). Institutionalization is considerably more common among seniors than among

younger people: overall, fewer than 2% of Canadians live in an institution, but at age 75 or older, the figure is 16%.[47]

In the absence of direct information in HPOI records about patients' place of residence, the postal code in combination with the PCCF+ and the Residential Care Facilities list was used to determine if patients lived in an institution. More than 20,000 institutional residents were identified and subsequently removed from HPOI using the PCCF+. Nonetheless, the coverage rate for seniors aged 75 or older remained significantly below the rates for younger people.

Use of the PCCF+ and the Residential Care Facilities list to identify institutions based only on the postal code is not ideal. Institutions that accounted for the majority of the population sharing a postal code had a higher chance of being identified and subsequently removed from the HPOI counts. As well, institutions in urban areas have more precise postal codes, and therefore, residents of such institutions were more likely to have been removed from HPOI. Rural and outlying suburban areas and smaller towns often have the same postal code for multiple enumeration/dissemination areas. Consequently, the coding is far less precise than for centralized urban postal codes, which are usually linked to a single enumeration/dissemination area. Therefore, residents of institutions in rural and outlying suburban areas and smaller towns likely remained in the HPOI counts.

The coverage rate in the territories is also problematic, in that the linked CCHS-HPOI estimates exceeded the standard (HPOI). This, however, is less of a concern, because the small number of CCHS records linking to HPOI (124) precludes future analyses featuring this subpopulation. Before the removal of on-reserve residents from the HPOI count, the coverage rate for the territories was 113%; after their removal, the rate was 163%. It is unclear why the linked HPOI-CCHS estimate is so much higher than HPOI. Records of CCHS respondents identified as living in the territories were reviewed to determine if some had high survey weights, which might explain the discrepancy between the HPOI and HPOI-CCHS counts. No discrepant weights were found; the average weight was 51, with weights ranging in value from 11 to 178.

In addition, hospitalizations pertaining to military personnel could not be identified and removed from HPOI. Full-time members of the Armed Forces are excluded from CCHS, and their inclusion may affect the coverage rate.

Conclusion

The value of record linkage is well established in epidemiological studies of population health. Linking information from routinely collected administrative health data such as HPOI with survey data like the CCCHS holds promise for discoveries about health determinants, different types of health care use and health outcomes. Coverage evaluation is a fundamental pre-requisite to analyses that integrate health-related information from multiple sources based on the CCHS-HPOI linked file.

This evaluation shows that the overall coverage rate is high, often over 90%, although some CCHS respondents, notably seniors, had lower rates. Even this limitation is manageable, however, as long as users of the file explicitly acknowledge that findings pertain only to the general household population (the target population of the CCHS), and not to the total population, particularly residents of institutions.

Acknowledgements

The author thanks Claude Nadeau for his assistance and thoughtful comments.

References

1. Fair ME, Whitridge P. *Tutorial on Record Linkage Slides Presentation.* 1997 Federal Committee on Statistical Methodology. Available at http://www.fcsm.gov/working-papers/RL_chap12.html.
2. Fair M. Generalized record linkage system-Statistics Canada's Record linkage software. *Austrian Journal of Statistics* 2004; 33(1 and 2): 37-53.
3. Villeneuve PJ, Morrison HI, Craig CL, et al. Physical activity, physical fitness, and risk of dying. *Epidemiology* 1998; 9(6): 626-31.
4. Weller I, Corey P. The impact of excluding non-leisure energy expenditure on the relation between physical activity and mortality in women. *Epidemiology* 1998; 9(6): 632-5.
5. Bata IR, Gregor RD, Eastwood BJ, et al. Trends in the incidence of acute myocardial infarction between 1984 and 1993 - The Halifax County MONICA Project. *Canadian Journal of Cardiology* 2000; 16(5): 589-95.
6. Roos NP, Havens B. Predictors of successful aging: a twelve-year study of Manitoba elderly. *American Journal of Public Health* 1991; 81(1): 63-8.
7. Khaw KT, Wareham N, Bingman S, et al. Combined impact of health behaviors and mortality in men and women: the EPIC-Norfolk prospective population study. *PLOS Medicine* 2008. Available at: http://medicine.plosjournals.org/perlserv/?request=get-document&10.1371/journal.pmed.0050012.
8. Weston TL, Aronson KJ, Howe GR, et al. Cancer mortality among males in relation to exposures assessed through a job-exposure matrix. *International Journal of Occupational and Environmental Health* 2000; 6(3): 194-202.

9. Aronson KJ, Howe GR, Carpenter M, et al. Surveillance of potential associations between occupations and causes of death in Canada, 1965-91. *Occupational and Environmental Medicine* 1999; 56(4): 256-69.
10. Wei L, Lang CC, Sullivan FM, et al. Impact on mortality following first acute myocardial infarction of distance between home and hospital: Cohort study. Available at: http://www.heart.bmj.com/onlinefirst.dtl.
11. Schmaltz HN, Southern D, Ghali WA, et al. Living alone, patient sex and mortality after acute myocardial infarction. *Journal of General Internal Medicine* 2007; 22(5): 572-8.
12. Ellison LF, Gibbons L, Canadian Cancer Survival Analysis Group. Five-year survival from prostate, breast, colorectal and lung cancer. *Health Reports* (Statistics Canada, Catalogue 82-003) 2001; 13(1): 23-34.
13. Ellison LF, Gibbons L. Leading cancers-changes in five-year relative survival. *Health Reports* (Statistics Canada, Catalogue 82-003) 2004; 15(2): 19-32.
14. Kaul P, Armstrong PW, Change W, et al. Long-term mortality of patients with acute myocardial infarction in the United States and Canada: Comparison of patients enrolled in global utilization of streptokinase and t-PA for occluded coronary arteries (GUSTO)-I. *Circulation* 2004; 110: 1754-60. Available at: http://circ.ahajournals.org/cgi/content/full/110/13/1754
15. Fischbacher CM, Bhopal R, Povey C, et al. Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: myocardial infarction in South Asians. *BMC Public Health* 2007; 7: 142
16. Jebamani LS, Burchill CA, Martens PJ. Links using data linkage to identify First Nations Manitobans: technical, ethical, and political issues. *Canadian Journal of Public Health* 2005; 96(Suppl.): S23-32.
17. Knoops KT, de Groot LC, Kromhout D, et al. Mediterranean diet, lifestyle factors, and 10-year mortality in elderly European men and women: the HALE project. *The Journal of the American Medical Association 2004;* 292: 1433-9.
18. Myint PK, Welch AA, Bingham SA et al. Fruit and vegetable consumption and self-reported functional health in men and women in the European Prospective Investigation into Cancer-Norfolk (EPIC-Norfold): a population-based cross-sectional study. *Public Health Nutrition 2007*; 10: 34-41.
19. Schnatter, A Robert. An analysis of death ascertainment and follow-up through Statistics Canada's Mortality Data Base System. *Canadian Journal of Public Health* 1990; 81(1): 60-5.
20. Schoenman JA, Sutton JP, Elixhauser A, et al. Understanding and enhancing the value of hospital discharge data. *Medical Care Research and Review* 2007; 64(4): 449-68.
21. Herrchen B, Gould JB, Nesbitt TS. Vital statistics linked birth/infant death and hospital discharge record linkage for epidemiological studies. *Computing Biomedical Research* 1997; 30(4): 290-305.
22. Chen J, Fair M, Wilkins R, et al. Maternal education and fetal and infant mortality in Quebec. *Health Reports* (Statistics Canada, Catalogue 82-003) 18; 10(2): 53-64.
23. Wen SW, Kramer MS, Liu S, et al. Infant mortality by gestational age and birth weight in Canadian provinces and territories, 1990-1994 births. *Chronic Diseases in Canada* 2000; 21(1): 14-22.
24. Kramer MS, Demissie K, Yang H, et al. The contribution of mild and moderate preterm birth to infant mortality. *Journal of the American Medical Association* 2000; 284(7): 843-9.
25. Bingham SA, Cassidy M, Cole TJ, et al. Validation of weighed records and other methods of dietary assessment using the 24h urine nitrogen technique and other biological markers. *British Journal of Nutrition* 1995; 73: 531-50.
26. Quigley MA, Hockley C, Davidson LL. Agreement between hospital records and maternal recall of mode of delivery: evidence from 12 391 deliveries in the UK Millennium Cohort Study. *International Journal of Obstetrics and Gynaecology* 2007; 114(2): 195-200.
27. Ford JB, Herrchen B, Gould JB, et al. Characteristics of unmatched maternal and baby records in linked birth and hospital discharge data. *Health Statistics Quarterly* 2007; 33: 25-33.
28. Fair M, Cyr M, Allen A, et al. *Validation Study for a Record Linkage of Births and Infant Deaths in Canada* (Statistics Canada, Catalogue 84-F0013) Ottawa: Statistics Canada, 1999.
29. Manuel DG, Schultz SE. Using linked data to calculate summary measures of population health: Health-adjusted life expectancy of people with Diabetes Mellitus. *Population Health Metrics* 2004. Available at: http://www.pophealthmetrics.com/content/2/1/4. Accessed: May 13, 2008.
30. Calver J, Brameld KJ, Preen DB, et al. High-cost users of hospital beds in Western Australia: a population-based record linkage study. *Medical Journal of Australia* 2006; 184(8): 393-7.
31. Houle C, Berthelot JM, David P, et al. Matching census database and Manitoba health care files. *Analytical Branch Studies Research Paper Series* (Statistics Canada, Catalogue 11F0019MPE) No. 91. Available at: http://www.fcsm.gov/working-papers/choule.pdf.
32. Brocco S, Visentin C, Fedeli U, et al. Monitoring the occurrence of diabetes mellitus and its major complications: the combined use of different administrative databases. *Cardiovascular Dialectology* 2007;15(6): 5.
33. Young TK, Kliewer E, Blanchard J, et al. Monitoring disease burden and preventive behavior with data linkage: cervical cancer among aboriginal people in Manitoba, Canada. *American Journal of Public Health* 2000; 90(9): 1466-8.
34. Ringland C, Correll PK, Lim K,et al. Hospital readmissions for asthma: a feasibility study comparing strategies for linking hospital morbidity data. *Australian and New Zealand Journal of Public Health* 2006; 30(5): 435-9.
35. Aronson KJ, Howe GR, Carpenter M, et al. *Occupational Surveillance in Canada: Cause-specific Mortality among Workers, 1965-1991* (Statistics Canada, Catalogue 84-546) Ottawa: Minister of Industry, 2000.
36. Nadeau C, Beaudet MP, Marion J. Deterministic and probabilistic record linkage. *Proceedings of Statistics Canada Symposium-Gatineau.* Ottawa: Minister of Industry, 2006.

37. Béland Y. Canadian Community Health Survey -methodological overview. *Health Reports* (Statistics Canada, Catalogue 82-003) 2002; 13(3): 9-14.

38. Statistics Canada. *Canadian Community Health Survey—Estimation.* Available at: www.statcan.ca/cgi-in/imdb/p2SV.pl?Function=getSurvey&SDDS=3226&lang=en&db=IMDB&dbg=f&adm=8&dis=2#b6. Accessed June 27, 2008

39. Statistics Canada, Household Surveys Methodology Division. External linkage production report: Data years: F1992 to F2004. (Unpublished), 2006.

40. Richards J, Brown A, Homan C. The data quality study of the Canadian Discharge Abstract Database. *Proceedings of Statistics Canada Symposium 2001. Achieving Data Quality in a Statistical Agency: A Methodological Perspective.* Available at: www.statcan.ca/english/freepub/11-522-XIE/2001001/session16/s16a.pdf. Accessed: July 2, 2008.

41. Statistics Canada. *2001 Census Dictionary* (Statistics Canada, Catalogue 92-378). Available at: www12.statcan.ca/english/census01/Products/reference/dict/appendices/92-378-XIE2002.pdf. Accessed June 26, 2008.

42. Statistics Canada. *Postal Code Conversion File (PCCF): Update.* Available at : www.statcan.ca/bsolc/english/bsolc?catno=92-153-UCB. Accessed June 28, 2008.

43. Statistics Canada. *Residential Care Facilities (RCF).* Available at: www.statcan.ca/bsolc/english/bsolc?catno=83-237-X. Accessed July 2, 2008

44. Fellegi IP, Sunter AB. A theory for record linkage. *Journal of the American Statistical Association* 1969; 64: 1183-210.

45. Rao JNK, Wu CFJ, Yue K. Some recent work on resampling methods for complex surveys. *Survey Methodology* (Statistics Canada, Catalogue 12-001) 1992; 18(2): 209-17.

46. Rust KF, Rao JNK. Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research* 1996; 5: 281-310.

47. Turcotte M, Schellenberg G. *A Portrait of Seniors in Canada* (Statistics Canada, Catalogue 89-519) Ottawa: Minister of Industry, 2007. Available at: www.statcan.ca/english/freepub/89-819-XIE/89-519-XIE2006001.pdf. Accessed January 4, 2008.

Appendix

Appendix Table A
Number and percentage of Canadian Community Health Survey respondents who agreed to have their survey responses linked with their administrative health records (HN) and who provided valid HN, by selected characteristics, Canada excluding Quebec, 2001

	Agreed to link		Agreed to link and HN valid	
	Number	%	Number	%
Total	**98,450**	**90.4**	**72,354**	**66.5**
Province/Territories				
Newfoundland and Labrador	3,533	91.3	2,933	75.8
Prince Edward Island	3,238	88.7	2,236	61.2
Nova Scotia	4,938	92.8	4,108	77.2
New Brunswick	4,634	92.8	3,746	75.0
Ontario	35,674	90.8	24,917	63.4
Manitoba	7,653	90.4	5,552	65.5
Saskatchewan	7,417	92.6	6,142	76.7
Alberta	12,757	88.2	9,155	63.3
British Columbia	16,493	90.1	11,990	65.5
Territories	2,113	83.9	1,575	62.6
Sex				
Female	52,865	90.5	40,334	69.1
Male	45,585	90.3	32,020	63.4
Age group				
12 to 24	19,246	91.8	13,538	64.5
25 to 34	14,482	90.9	10,119	63.5
35 to 44	18,892	90.2	13,883	66.3
45 to 54	16,036	89.6	11,905	66.5
55 to 64	11,493	90.0	8,695	68.1
65 to 74	9,778	90.3	7,684	71.0
12 to 74	89,927	90.5	65,824	66.3
75 or older	8,523	89.5	6,530	68.6

Source: 2000/2001 Canadian Community Health Survey